DOCUMENT RESUME

ED 176 216 CS 004 962

AUTHOR Kean, Michael H.; And Others
TITLE What Works in Reading? Summary and Results of a Joint School District/Federal Reserve Bank Empirical Study in Philadelphia.
INSTITUTION Philadelphia School District, Pa. Office of Research and Evaluation.
PUB DATE May 79
NOTE 67p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Class Size; Elementary Education; Grade 4; *Grouping (Instructional Purposes); Principals; *Reading Achievement; Reading Consultants; *Reading Instruction; Reading Materials; *Reading Programs; *Reading Research; Teaching Techniques
IDENTIFIERS *Basal Reading; *Reading Teachers

ABSTRACT

More than 1,800 Philadelphia fourth grade students were involved in a study of reading instruction in the Philadelphia elementary schools conducted during the school year 1975-76. Four critical factors were noted as improving reading achievement in the district's elementary schools: the use of the linguistic basal approach to reading instruction, former reading professionals as principals, more time spent in the classroom during class hours by the teacher as opposed to professional or nonprofessional absences, and the use of a combination of whole class instruction and small group instruction. Some of the variables that produced no difference in reading achievement scores were the mobility of the students and busing, the amount of education and experience of the principal, the number of graduate courses in reading taken by the classroom teacher, the professional preparation and experience of the school reading teacher, the time spent on reading instruction, and the small size of classes. (MKM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

- ED176216

WHAT WORKS IN READING?

Summary of a Joint School District/ Federal Reserve Bank Empirical Study in Philadelphia

Michael H. Kean
Anita A. Summers
Mark J. Raivetz
Irvin J. Farber

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

The School District of Philadelphia

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)"

CS004962

May 1979

Office of Research and Evaluation
The School District of Philadelphia
21st Street and The Parkway
Philadelphia, Pennsylvania 19103

WHAT WORKS IN READING?
Summary and Results of a Joint School District/ Federal Reserve Bank Empirical Study in Philadelphia.

Michael H. Kean
Anita A. Summers
Mark J. Raivetz
Irvin J. Farber

May 1979

Office of Research and Evaluation
The School District of Philadelphia
21st Street and The Parkway
Philadelphia, Pennsylvania 19103

WHAT WORKS IN READING?

A. BACKGROUND TO THE STUDY

The purpose of this study was to determine what makes a difference in teaching children to read. The years of experience of many thousands of teachers have testified to the fact that no one, simple, magical solution could be expected. It was, therefore, the task of this study to try to identify the factors that contributed most strongly to achievement growth in reading. In particular, there was interest in appraising the relative merits of approaches to reading.

In 1975, the Federal Reserve Bank of Philadelphia published a study of achievement growth using sophisticated statistical procedures; among them multiple regression. Though these procedures are quite commonly used in the field of economics, they were relatively rare in education.

In 1973, the Office of Research and Evaluation initiated several studies which attempted to identify the practices and characteristics of schools with successful reading programs. Over a period of several years, more than eighty characteristics were examined, but the results were inconclusive because the characteristics did not seem to differentiate the more successful from the less successful schools.

In October 1975, the Superintendent of Schools and the President of the Federal Reserve Bank of Philadelphia were invited by the city's Deputy Mayor to meet with him to discuss the Federal Reserve Study. When it became evident that the School District had no intention of using the results of the 1975 study, it was suggested that a new study be conducted utilizing the sophisticated statistical techniques of the old study, but attending to the School District's primary concerns. This study resulted from that agreement.

B. HOW THE STUDY WAS CONDUCTED

Schools were picked for the study by rank-ordering all elementary schools by their overall average reading scores (1975, grades 1-4) and their overall gain scores (1974-1975, grades 1-4) in reading. Ten schools at the top of both lists, and ten at the bottom of both lists were selected. Five schools in the middle of both lists completed the sample of 25 schools.

Though pupils in this study were from the 25 selected schools, the study involved facts about individual pupils. It was the individual pupil who was the subject of this study. All of the fourth grade pupils in the schools selected were in the study. Fourth grade was chosen because it is the grade in which a large drop in test

scores is typically observed; it is the highest grade which is found in all elementary schools; and the ability to read at that grade is importantly related to a students subsequent progress.

Information was gathered for the study by teams made up of the staff of the Office of Research and Evaluation and of the Division of English/Reading Language Arts Education. These teams visited the schools involved and interviewed the principal, the teachers, the reading teachers, and the reading aides. A total of 25 principals, 25 reading teachers, 94 classroom teachers and 68 reading aides were interviewed. In addition, team members gathered large amounts of information from each pupil's records.

All of this information was checked and double checked for accuracy and then placed on a computer file. One hundred sixty-two facts (or variables) about each of 1,828 pupils made up the computer file on which the study was based. Later, various combinations of these separate items of information increased the total number of variables about each pupil to 245.

Five kinds of facts about each pupil were gathered:

1. Facts about the principal of the pupil's school.
2. Facts about the reading teacher in the pupil's school.
3. Facts about the pupil's classroom teacher.
4. Facts about the pupil's school.
5. Facts about the individual pupil.

Once all of the information about each pupil was placed on computer tape, a random sample of 25% of the pupils was removed from the file and put aside. This was kept as a cross-validation sample, which allowed the results to be double checked.

Over 500 multiple regression equations were run on the large (75%) file. This involved trying out many different combinations of facts (equations) about the pupils in order to determine which combination was most closely related to gains in reading scores. After the equation which seemed to be most closely related to gain in reading scores was found, it was used on the cross-validation (25%) file. Findings that were the same in both cases could be reported with some confidence; those that were not could be presented for discussion and further study.

C. THINGS THAT MADE A DIFFERENCE

Based on these procedures, the following was found:

1. Some things about the pupil seemed to make a difference:

   a. The better the pupil's attendance, the more the reading score increased.

   b. The fewer unexcused absences, the more the reading score increased.

   c. Pupils who attended kindergarten seemed to gain more than those who did not.

2. Some things about the principal seemed to make a difference:

   a. Where principals had experience in the field of reading, the pupils achieved better.

   b. The more the principal observed in the classroom, the better the pupils achieved.

3. Some things about the school seemed to make a difference:

   a. When schools had more professional supportive staff per pupil, the pupils achieved better.

   b. Pupils clearly achieved worst in K-6 schools; K-7 schools were best; K-8 and K-5 schools were in between.

4. Some things about the classroom teacher seemed to make a difference:

   a. Where teachers had more pay periods without absences, the pupils achieved better.

   b. Where teachers had more than one year experience teaching fourth grade, the pupils achieved better.

5. Some things about reading instruction seemed to make a difference:

   a. Pupils being taught using the linguistic basal approach achieved distinctly better than pupils using other reading approaches.

   b. The more minutes a week of sustained silent reading, the better the pupils achieved.

   c. The fewer reading aides per week, the better the pupils achieved.

   d. Where teachers felt that, if given a choice, they would pick the same reading program again, the pupils achieved better.

   e. Pupils in larger classes (up to 35) gained more than those in smaller classes (as few as 24).

f. Pupils in classes taught in a small group/whole class combination achieved better than pupils who were taught individually only, in small groups only, or as a whole class only.

D. THINGS THAT DID NOT MAKE A DIFFERENCE

Though a number of things were found to make a difference, most did not. A complete list will be found in the full study, but some things found not to make a difference are noted here because, in the past, they have been considered important. It was found that the following things did <u>not</u> affect pupil achievement:

1. The number of times a pupil moved since starting school.
2. Being bussed to relieve overcrowding.
3. The amount of administrative experience the principal had.
4. Whether or not the principal had a doctorate.
5. The number of years the principal had been in the school.
6. Whether the pupil came from a lower or higher income neighborhood.
7. Whether the school had more or fewer pupils from lower income families.
8. The race of the teacher.
9. The number of years of experience of the teacher.
10. The number of graduate courses in reading and language arts taken by the teacher.
11. The number of minutes a day of reading instruction.

12. The time of day reading was taught.

13. The professional preparation and experience of the school reading teacher.

E. A DEEPER LOOK AT THE DATA

At times, certain methods or problems are thought to affect higher achievers more than lower achievers, or vice versa. For that reason an attempt was made to determine if this was the case with a number of the facts that were studied. Some of the more interesting findings were:

1. The lowest achieving pupils did better if they were in higher achieving schools.

2. The linguistic basal reading approach was more beneficial to middle and higher achieving pupils than other reading approaches, but the reading approach utilized did not make any difference for the lowest achieving pupils.

3. Unexcused absences were equally harmful to higher and lower achieving pupils.

4. Having attended kindergarten was equally beneficial to higher and lower achieving pupils.

F. THINGS TO KEEP IN MIND

In reviewing these findings and deciding what they mean, it would seem important to consider the following:

1. This study was conducted only on fourth graders. It is not likely that fourth graders are completely different from all other grades, but one should be careful in thinking that all of the findings would be exactly the same for first graders or twelfth graders, for example.

2. Educational programs in schools are ever changing. This study was conducted during the fifth year of the District Reading Programs. Some of the things found to be of no importance at that point might possibly have been of critical importance during the first year or two of the program.

3. This study suggested what combination of things made a difference in reading achievement. It did not tell why. To try to find out why, additional studies of a different kind would be needed.

4. As conditions change, as the reading program changes, and as suggestions from this study are tried out in the schools, it would be important to repeat this study, in order to find out something about the impact of these changes on reading achievement.

G. SOME POLICY IMPLICATIONS OF THIS STUDY

Based on the findings of this study, it is recommended that the following action be considered:

1. Establish an outreach program to assure that as many children as possible have a kindergarten experience.

2. Explore the expansion of the use of the linguistic basal reading approach.

3. Consider the establishment of a special task force to explore programs for, and develop intervention studies of, the lowest achieving pupils.

4. Develop or modify programs or policies to minimize pupil and teacher absenteeism.

5. Develop or modify programs or policies to increase the amount of teacher time that is spent in direct instructional contact with pupils.

6. Develop or modify programs or policies to increase the amount of time principals spend observing in classrooms.

7. Explore the development of programs to increase the amount of knowledge and experience in reading which potential principals have, prior to their appointments.

8. Explore reasons why the K-7 organization seemed to be associated with the greatest amount of pupil gain, and K-6 the least.

9. Consider mandating a policy of daily sustained silent reading in all schools.

10. Initiate discussions with appropriate governmental officials directed toward the modification of current regulations to permit the allocation of federal resources on a pupil specific basis.

H. A FINAL WORD

A final word of caution seems in order. The findings presented are the ones which seem most pertinent, and on which it is possible for the school system to take action. Complete information on the entire study, and a list of all variables, can be found in the full report. The findings are the result of intensive investigation using sophisticated mathematical techniques. As such, they should certainly be given serious consideration and thought. There is, however, no guarantee that implementation of recommendations based on these findings would cause any miracles. There are too many things involved to make any such guarantee. These findings are, however, the best information available at this time. Even without guarantees, it is better to be guided by such information than to try to make decisions in a vacuum.

WHAT WORKS IN READING?

The Results of a Joint School District/ Federal Reserve Bank Empirical Study in Philadelphia

Michael H. Kean
Anita A. Summers
Mark J. Raivetz
Irvin J. Farber

May 1979

Office of Research and Evaluation
The School District of Philadelphia
21st Street and the Parkway
Philadelphia, Pennsylvania 19103

THE SCHOOL DISTRICT OF PHILADELPHIA

Superintendent of Schools
Dr. Michael P. Marcase

Deputy Superintendent of Schools
Mr. Charles A. Highsmith

BOARD OF EDUCATION

Mr. Arthur W. Thomas, President
Mrs. Edward Oberholtzer, Vice President
Mr. Augustus Baxter
Mrs. Lawrence Boonin
Mr. Samuel H. Rubin
Robert M. Sebastian, Esq.
Mrs. Michael Stack
George Philip Stahl, Jr., Esq.
Dr. Nicholas E. Trolio

Office of Research and Evaluation
Dr. Michael H. Kean
Executive Director

TABLE OF CONTENTS

Acknowledgements		ii
The Study		
I	Background	1
II	The Rationale for the Study	4
III	What Preceding Studies Indicate	7
IV	Procedures for Data Collection and Organization	12
V	Model and Estimation Procedures	25
VI	What Works in Reading?	30
VII	Policy Implications	38
References		45
Sources of Data		47
Digest of Results		48

ACKNOWLEDGEMENTS

The authors wish to express their appreciation to the numerous people who provided assistance, encouragement and support to this study.

Michael P. Marcase, Superintendent of Schools

Bernard G. Kelner, Associate Superintendent for Curriculum and Instruction

Marjorie Farmer, Executive Director for Reading/English Language Arts

Barbara L. Wolfe, Assistant Professor, Department of Economics, and Preventive Medicine, and Research Affiliate, Institute for Research on Poverty, University of Wisconsin, Madison

Ellery M. Pierson, Manager for Systems Development, Data Management and Production, Office of Research and Evaluation

Daniel Ashler, Research Associate, Office of Research and Evaluation

Randy Stratt, Summer Intern, George Washington University School of Law

Staff members from the Office of Reading/English Language Arts

Staff members from the Office of Research and Evaluation

The authors express their thanks to the principals, reading teachers, classroom teachers and reading aides who were gracious enough to give of their time, thereby making this study possible.

To the many others who have given of themselves throughout this study, the authors offer their sincere appreciation.

I. BACKGROUND

The study had its genesis in two independent efforts by the School District of Philadelphia and the Federal Reserve Bank of Philadelphia.

As part of its ongoing research and evaluation activities, the Office of Research and Evaluation of the School District of Philadelphia had conducted several studies of reading practices in recent years. One study--that of a single school which had demonstrated remarkable progress in reading during the course of one school year--was a post hoc attempt to describe the reading effort in as much detail as possible.[1] It was largely "a fishing expedition." The results were interesting, but served primarily as an impetus for a more intensive look at successful reading practices.

A more extensive study, involving fifteen "successful" schools was conducted the following year (1974).[2] This study involved interviews of administrators and students, as well as the examination of a number of school characteristics. Though over eighty variables measuring these characteristics were examined, the results were again inconclusive. Variables could not be identified that differentiated the schools studied from the other "less successful" schools in the system.

At the same time, the Federal Reserve Bank of Philadelphia was conducting a study of its own. In 1975 it published a study of achievement growth in the Philadelphia schools, which utilized econometric techniques.[3]

---

[1]Scheib, J.E. and Comerford, J.P. Reading Improvement at Vaux Jr. High School, Office of Research and Evaluation, School District of Philadelphia. 1972.

[2]Office of Research and Evaluation. Successful Schools Study. 1974. (unpublished)

[3]Summers, A.A. and Wolfe, B.L. "Which School Resources Help Learning? Efficiency and Equity in Philadelphia Public Schools." Federal Reserve Bank of Philadelphia Business Review. February, 1975.

This effort was widely acclaimed, especially by economists, but was criticized from within the School District of Philadelphia. Objections were raised on both a technical and on an emotional level. After a period of heated interaction, it became evident that the School District had no intention of utilizing this study for policy development or decision making purposes.

In late October of 1975, the Superintendent of Schools of Philadelphia, and the President of the Federal Reserve Bank of Philadelphia were invited by the Deputy Mayor of Philadelphia to meet with him for the purpose of discussing the recently released Federal Reserve Study. Both brought members of their staffs to the meeting. The general focus of the discussion that morning was what, if any, followup to the findings and recommendations of the Federal Reserve Bank Report was being planned. It was suggested that, in spite of School District comments critical of certain of the technical procedures of the report, there was a considerable amount of promise in the methodology and implications for longer range policy decisions.

An agreement was reached whereby the Office of Research and Evaluation of the School District and the Research Department of the Federal Reserve Bank would meet to discuss the development of a followup study utilizing, as a base, the methodology employed by the Federal Reserve Bank, but integrating within it the concerns and learning priorities of the School District. Shortly after this meeting at City Hall, Michael H. Kean, the Executive Director of the Office of Research and Evaluation, and Anita A. Summers, Research Officer and Economist of the Federal Reserve Bank met to discuss the possibilities for a new study.

Key staff members from both the Federal Reserve Bank and the School

District met in a series of half day work sessions to plan the study. At the outset the tone was cordial but guarded, and the meetings were negotiating sessions as much as planning meetings.

Reasons for earlier disagreement quickly became evident. The nature of various school variables had to be clarified. In addition, differences in statistical terminology were a barrier. At one point, a staff member familiar with both multiple regression analysis and analysis of variance was brought in to "translate" the terms and concepts used.

As the work sessions progressed, they brought with them a heightened awareness of each other's world. Eventually a genuine mutual respect developed. While this commentary might seem extraneous to the proper purpose of this paper, the understanding and regard referred to above were sine qua non for the implementation of this study.

II THE RATIONALE FOR THE STUDY

Some schools are more successful than others in terms of reading achievement growth from one year to the next, and in terms of annual level of achievement. What are the reasons that one school has a higher score than another? Why do some students perform better than others? By collecting measures of possible reasons, one might be able to discriminate among reasons by virtue of their degree of association with the measure of success. In short, it is possible to compute the multiple regression between the achievement measure (the dependent variable) and the estimated measures of the hypothesized reasons (the predictors or independent variables).

This study is a search for those variables that are associated with changes in achievement growth, and an attempt to describe the associations in terms of socio-economic conditions, educational inputs, and peer group characteristics. The cliche "correlation does not mean causation" is an appropriate caveat. However, in the world of educational policy, decisions must be made on the basis of the best available information. Therefore, policy implications are the end product of this study.

During the 1970-71 school year, the School District of Philadelphia initiated an intensive thrust in the area of reading instruction. This coincided with the national "Right to Read" emphasis. As a means of implementing this effort, responsibility was placed directly in the offices of the superintendents of the eight administrative subdistricts comprising the School District of Philadelphia. These districts ranged in size from seventeen thousand to forty thousand students.

Resources were made available to the district superintendents (both

Title I and operating budget funds were included), and each district was directed to produce reading plans detailing its programs and how the resources provided were utilized.

The result of this pattern of organization was the production of eight different reading plans, each reflecting the pupil population, the available resources, and the personnel of the district. Moreover, the degree of uniformity within each district varied widely from plans based largely upon individual school option to plans mandating specific materials for specific target groups. At one time, in fact, there were thirty-two different reading programs being utilized within the School District.

From the first year of the program, the Office of Research and Evaluation published a catalog of reading programs. In constructing this catalog, it became clear very quickly that many differences were more apparent than real. Some terms meant different things in different districts, while in other districts, differing terms were used to identify the same things. To attempt to remedy this situation, the Office of Research and Evaluation and the Office of Reading/English and Language Arts combined efforts to produce a taxonomy of reading "approaches" and a glossary of reading "programs." The intent of this effort was to enhance communication by imposing a common terminology. Essentially, an "approach" was defined as that philosophical concept on which programs are based (i.e. what each program assumes about how reading is learned), while a "program" was defined as that set of commercial (or local) materials used as a curricular base.

Program evaluation was conducted on a district-by-district basis. While this level of evaluation service was helpful to individual districts, it was of limited value for systemwide decision-making. It was still necessary for the school district to identify the pattern (or patterns) of reading instruction associated with the greatest improvement in reading achievement. Only in this way would it be possible to provide schools and subdistricts with information about less successful practices and recommend the institution (or maximization) of more successful ones.

The present study, then, is an attempt to generate information about those attributes and practices associated with the greatest improvement in reading achievement -- and to generate them in a way that would be useable for planning long range school district policy in the area of reading instruction.

III WHAT PRECEDING STUDIES INDICATE

To review all of the literature that relates to "what works" in reading, and that relates to the problems of large scale empirical studies in education would take volumes. Some comments on how these writings relate to the original 1975 Fed Study, and their relevance to this new cooperative study between the Federal Reserve Bank of Philadelphia and the School District of Philadelphia seems appropriate.

A repeated finding of many studies searching for answers to questions about educational achievement has been that the student's background largely determines his or her school performance -- family income and race, particularly, are the most important determinants. The findings of Coleman (1966), Jensen (1969) and Jencks (1972) suggested that schools were relatively ineffective tools for promoting the achievement growth of students. Averch (1972) concluded that "the best information we have . . . is that schools do not now have a tremendous impact on the achievement that does occur."

A. The 1975 Federal Reserve Bank Study

Much has happened since 1972. Educators are no longer forced to react to the notion that "schools make no difference." A contribution to this was the publication of the original Federal Reserve Bank study.

The February 1975 Federal Reserve Bank of Philadelphia's Business Review published the Summers and Wolfe study "Which School Resources Help Learning? Efficiency and Equity in Philadelphia Public Schools." The study concluded that certain school inputs do make a difference in achievement growth. To summarize the major findings (for elemementary schools):

.Students experienced larger rates of growth in achievement if they attended more, and their unexcused absences and latenesses are minimized.

.Students do better if their teachers graduated from higher-rated colleges.

.Students do better if their schools' racial make-up is 40-60 percent Black; if they are in classes of 33 or less; and if they are in schools with more high achievers.

.Black students perform better in smaller elementary schools.

.Low achieving students do better in classes with less than 28 students; and if they have new teachers.

.High achievers do best with experienced teachers.

The Federal Reserve Bank broke new ground by examining educational inputs using a large amount of pupil-specific data. Multiple regression analysis was used to analyze the data. Equally important, the data allowed examination of the relationship between school variables and pupil characteristics; that is, it allowed examination of whether or not inputs had different effects on different sorts of students. Summers and Wolfe did not just ask statistically whether being in a small class was better, for example, but whether or not it was better for high achievers, medium achievers, and low achievers to be in smaller classes.

The Federal Reserve Bank study found that the effects of many inputs were very different, and that the lack of importance assigned to school inputs in previous studies was likely to be due to the absence of pupil specific data; and associated with that, the absence of examining the differential impacts of the inputs. Other studies failed to uncover some of the effective accomplishments of schools because many school inputs did not consistently benefit all students. If, for example, low-income students require a different proportion of individualized instruction than high-income students, and the researcher looks only at

the average amount of individualized instruction in the school, a finding that individualized instruction has no effect would disguise its true impact. Summers and Wolfe (1977) concluded that the reason educational studies have failed to find the things schools do are effective is the limited data which were specifically tied to the pupil. Several reviews of the literature on school effectiveness, most notably Barr and Dreeben (1977) make this point most clearly.

Burstein (1978) provides further support for the current study. In citing Hanushek (1970,1972); Murnane (1975); Summers and Wolfe (1974); and Winkler (1972); he concludes that school effects studies using the student as the unit of analysis are more likely to yield accurate estimates of the factors influencing individual student achievement.

B. Explanatory Observational Studies

The use of a production function model to examine school effects is, by now, commonly accepted. In this concept, taken from economics, the school is treated as transforming inputs (e.g., books and equipment, teachers time, programs) into outputs (gains in achievement). Reviews on the major contributions of production function approaches to school effects have been completed by Averch (1972); Spady (1973); and Alexander and McDill (1976). They also provide support for the procedures involved in the current study.

The current study, as well as its immediate predecessors, might best be described as "explanatory observational studies." Such efforts are discussed extensively by Cooley (1978). These are "studies of such phenomena as program effectiveness, in which an attempt is made to explain why the students in some programs achieve more than students in others."

Cooley highlights the importance of giving more attention to developing methods for conducting such studies.

The importance of Cooley's work is his support for such studies. Lack of randomness, lack of controlled treatment, and lack of independence are discussed as reasons why many observational studies are criticized.

Cooley discusses the procedures for meeting these criticisms under three topics; the sampling framework, the theoretical model, and the statistical procedure. He then develops requirements for sound observational studies that "can yield consistent, convincing, useful explanations of educational phenomena." The Philadelphia School District/Federal Reserve Bank study attempts to meet Cooley's requirements without being apologetic for not conducting an intervention study.

C. Reading Studies

While not new, Chall's (1967) study, Learning to Read: The Great Debate is still among the best at sorting out the research on "what works" in reading.

Chall presents several major observations and conclusions. They include: (1) methods which emphasize the sound value of letters (phonics/decoding) produce better reading achievement than methods with a meaning emphasis; (2) the child reads faster in the early stages of learning through a meaning emphasis, but by fourth grade the advantage is lost, probably due to poor word recognition; and (3) children of average or below average scholastic aptitude probably perform better under code emphases.

A review of other more recent studies -- Guthrie et al (1976); Della-Piana and Endo (1973), Edmunds and Federiksen (1977), suggest support of Chall's findings.

D. Summary

Learning reading skills is a difficult task that takes place in an extremely complex environment. The major strength of this study is that extensive pupil specific data on the reading approach; on the way in which reading instruction was delivered; and on teacher, school, and socio-economic characteristics were available. This study attempts to incorporate the recent findings on statistical techniques for exploring educational achievements and findings on reading in an extensive, systematic look at fourth graders. It attempts to capture more of "what goes on in the classroom" than has been done before.

IV PROCEDURES FOR DATA COLLECTION AND ORGANIZATION

A. Selecting the Data Set

The process of deciding what to include in the study was an exacting and time-consuming one. Possible variables were nominated for consideration based upon the collective wisdom of past studies, the practical suggestions of reading specialists, and the experience of the Office of Research and Evaluation of the School District of Philadelphia. Each potential variable was then examined in the light of three considerations:

- Information Considerations
- Logistical Considerations
- People Considerations

Only those variables which passed all three "tests" were identified as study variables. Briefly, the considerations were defined as follows:

1. Information Considerations

   The first task was to justify the need for each item of information. What might it tell? What action could be taken on the basis of it? What was the reason for believing that it was of importance? What might be the consequences of omitting it?

2. Logistical Considerations

   It was necessary to consider where the information existed and what would be involved in retrieving it. Wherever possible, unobtrusive approaches were used. If reliable records already existed, personal contacts could be minimized.

   It was necessary to estimate the staff resources needed to gather each item of information and to judge the feasibility of doing so in light of the time and the personnel available.

The fact that the end of the school year was rapidly approaching was in itself an important consideration, and dictated the use of as many data collectors as possible in the shortest possible time span.

Collection instruments that would permit efficient retrieval of information from records had to be designed. Questionnaires had to be short enough to permit administration in not more than 15 minutes. This was a practical consideration; experience indicated that the amount of cooperation that could be expected from schools, all other variables being equal, was inversely related to the amount of disruption caused.

3. People Considerations

This was a two-sided problem. A great deal of status information would have to be requested of teachers and administrators, (i.e., personal status, and operational status of the program.) Questions had to be framed that individuals would be willing to answer, as well as able to answer. Questions could not be worded so that the only possible answers were the socially acceptable ones. Beyond this, it was necessary to select the "correct" person to whom to address a question. For example, one could not very readily ask an individual whether he was performing an activitity, if it was clear policy that he do so. In essence, one would be asking whether or not someone was doing his job. It would not, of course, be possible to take at face value the answer to such a question. Instead, it was necessary to address such questions to individuals who were to receive the service, so that such conflicts could be avoided.

The second side of the problem involved the interviewers. It was desirable to design instruments involving a minimum amount of interviewer judgement and interpretation.

While such considerations might not weigh heavily upon laboratory studies, they are of critical importance in field studies. Failure to give proper attention to them could likely result in the generation of data in which the researchers had little confidence.

B. Design of Instruments

Once the list of independent variables was developed, five data collection instruments were designed. One was merely a form to permit efficient recording of pupil specific data from pupil records. The other four were interview protocols. Questions in each were clustered into several broad areas in order to enable the interviews to flow more easily.

1. The Principal Interview

This protocol included two broad areas of questions: (a) the principal's professional development and (b) the principal's management of the school reading program.

2. The Reading Teacher Interview

Three broad areas of questions were included in the Reading Teacher Questionnaire: (a) teacher's professional development; (b) school district support climate (i.e., services received by the reading teacher); and (c) reading teacher's services to the school.

3. Classroom Reading Aide Interview

Only four questions were contained in this questionnaire. They related to the background of the aide and the services provided.

4. Classroom Teacher Interview

This was the longest interview and included questions in the following areas: (a) teacher's professional development; (b) school district support climate (i.e., services received by the teacher); (c) classroom descriptions (i.e., classroom organization and practice), and (d) reading materials used.

When the first drafts of the instruments were completed, they were pilot tested in several schools which would not be included in the final sample. Revisions were made, especially where questions were found to be confusing to interviewers or where questions invited too strongly the more socially desirable response.

C. Selecting the Study Population

The later work sessions involved the selection of the study population. It was agreed that though the study was to be conducted on individual pupils, selection was to be by school. This made it possible to take into consideration two aspects of achievement--absolute level and growth.

The following procedure was used to determine the schools from which the study sample was drawn:

1. The 1974 average Total Reading Achievement Development Scale Score (ADSS) on the California Achievement

Tests (CAT-70) were summed over grades 1 to 4 for all 190 elementary schools in the District.

2. Step 1 was repeated with the 1975 average Reading Total ADSS CAT-70 scores.
3. Schools were then ranked on the difference of these sums and schools ranked 1-15, 89-103, and 176-190 were identified.
4. The schools were also ranked on the February 1975 CAT Total Reading ADSS for grades 1-4 and the schools ranked 1-15, 89-103, and 176-190 were identified.
5. The dual rankings were examined for "doubles." Two schools appeared in both high-achievement and high-gain groups, and two schools appeared in both low-achievement and low-gain groups. These four schools were the first selected for the study. "High achieving" schools were selected if they were in the top quarter in achievement and top third in gain; "middle" schools were selected if they were in both the achievement and gain "middle" groups; "low" schools were selected if they were in the bottom quarter on achievement and bottom third in gain.
6. Schools were selected to provide representation from all eight administrative sub-districts of the city.
7. The final sample contained ten "high-high," 5 "middle-middle", and 10 "low-low" schools for a total of 25.

The 25 schools included in the study were representative of students in schools having high, middle, and low degrees of success in reading achievement on the two dimensions described and also provided for

the inclusion of a geographical cross section of the city. All fourth graders in these schools became subjects in the study.

Grade four students were selected for study for several reasons. First, it is an important point at which to identify trouble: a sharp decline in test scores at this grade has been observed over a number of years. Second, it is the highest grade that would permit all elementary school pupils to be potential subjects. Since Philadelphia elementary schools are organized in a number of grade configurations (i.e., K-4, K-5, K-6, K-7, K-8), selection of a higher grade would have eliminated some school organizations, thereby eliminating the ability to examine the impact of school organization on learning. Finally, the grade three score was selected as the pretest score because it was thought necessary to avoid the Level I CAT-70 Test (used in grades 1 and 2) due to problems in the norms tables (i.e., the 50th percentile in Grade One comprehension is one item above guessing level.) Both Grades three and four are given the Level II Test, so that both their achievement levels and their growth scores could be used.

It should be noted that while subjects were selected from schools utilizing 1975 scores and the changes from 1974 to 1975, the study was conducted using individual students' 1976 test scores and the change from 1975 to 1976. Also, though schools were selected using the Total Reading score, the dependent variable was the student's score on the Comprehension subtest. This decision, too, was taken to avoid problems of the insensitivity of the norms tables to students scoring above the 50th percentile in the Vocabulary subtest.

D. Collecting the Data

Once the sample was selected, and the final interview and data collection devices were complete, teams composed of staff from the Office of Research and Evaluation and reading program staff from central and district offices were assigned to the twenty-five schools in the study. Schools were, of course, given advance notice, and the letter explaining the data collection effort was countersigned by the Superintendent of Schools. The Philadelphia Federation of Teachers was kept informed of the effort. An orientation session was held at which staff was oriented to the purpose of the study and the procedures and instruments to be used. Teams of two or three persons were then sent to each school involved.

Depending on the size of the school, teams spent from one-half day to two days interviewing school personnel and recording data from pupil records. In all cases, principals and classroom teachers were interviewed by the Office of Research and Evaluation personnel. The task of interviewing the reading aide and the reading teacher was shared with personnel from the Division of Reading. In some cases, team members covered classes so that teachers could be interviewed. Information concerning pupils was copied from school records on the specially developed recording forms.

Administrators and teachers were generally very cooperative, thus enabling the entire data collection process to be completed in less than two weeks. Within the time-frame, twenty-five principals, twenty-five reading teachers, ninety-four classroom teachers, and sixty-eight reading aides were interviewed. Records of about two thousand children were examined and predetermined information recorded.

E. Organizing the Data

The data collection effort resulted in a tremendous amount of information in need of processing. This mass of data caused a somewhat pleasant problem, however. Information on 25 schools, 25 principals, 25 reading teachers, 94 classroom teachers, 68 reading aides and 1,828 fourth grade students provided an incredibly rich data base.

Questionnaires were examined for completeness, reduced and coded where necessary, and examined again. Not only were Principal's Questionnaires compared with one another, but all four types of questionnaires from each school were examined to check for consistency. (This procedure led to the elimination of the Aide Questionnaire. Since an aide could not be attached to a student with 100% accuracy, and the student was the unit of observation, only aide service, "yes or no", became a variable assigned to a student's classroom.)

The within-school questionnaire comparisons also led to the elimination of several "control" questions included on the instruments for principal, reading teacher, and classroom teacher. These questions were related to reading program planning and service. There was sufficient consistency among the respondents to verify the accuracy of the answers. In all, some 11 questions were used for verification purposes only: five from the Reading Teacher Questionnaire, four from the Classroom Teacher Instrument, and two from the Principal Questionnaire.

Separate files, each with frequency distributions, were created from the principal, reading teacher, classroom teacher, school and student data. These files were sorted and merged into the final master file.

The master file held a record for each student containing all the variables that applied to the student as a result of classroom and school assignment.

Careful quality control procedures throughout the data collection phase of the study provided every reason to believe that the data collected were accurate, and warranted extensive statistical analysis.

F. Choice of Variables

1. The Dependent Variable

This was the topic of much discussion throughout the study's planning phase. For each student there existed two achievement measures on the Reading Comprehension[4] subtest of the California Achievement Test, 1970 edition (CAT-70)--the February 1975 (third grade) and February 1976 (fourth grade) scores. Discussion centered on whether the fourth grade score should be the dependent variable (with the third grade score as one of the independent variables), or whether the difference between the third and fourth grade score should serve as the dependent variable, (with the third grade score as one of the independent variables). In short, should the dependent variable be an achievement score or a growth/gain score? The latter was selected, since it is the impact of input changes on reading development that is the focal issue.

The CAT-70 presents scores in terms of Achievement Development Scale Score (ADSS) Units, thereby avoiding the problems of subtracting

---

[4]The selection of the Reading Comprehension subtest as an achievement measure is discussed more fully in the preceding section.

grade equivalents or percentile ranks. The Dependent Variable was defined as the growth score from third to fourth grade in Achievement Development Scale Score Units for Reading Comprehension on the CAT.

2. The Independent Variables

The data collection effort yielded 162 separate variables. Dummy variables and interaction variables[5] eventually increased this number to 245.

The original files included Principal, Reading Teacher, Classroom Teacher, School, and Student variables. By assigning the initial letter of the file to the first character of the label, variables were immediately identifiable:

.P (Principal)

.R (Reading Teacher)

.T (Classroom Teacher)

.S (School)

.K (Kids/Students)

The remaining seven characters were abbreviated descriptions of what the variables measured. As a result, a rather strange language developed. Discussions concerning TNAIDEWK (Teacher variable; number of times the teacher has aide service per week), KDAYPRES (Kid variable; number of days present), or SGRDORG6 (School variable, grade organization K-6) were not uncommon.

[5]Dummy variables and interaction variables may require some definition. In order to analyze the data, variables are recorded as numbers. In the case of student absences, for instance, there is no problem. Dummy variables are used when recording nominal data such as sex, race, or school organization. Numbers are assigned for identification purposes: for example males are 1, females 2. An Interaction variable is the product of two independent variables. Multiplying a student's third grade score by the number of times absent, for example, creates a new interaction variable which is used to test whether or not student attendance makes more or less difference if the student is a high, middle or low achiever in reading. The impact of absences on a student is measured by combining the coefficients of the absence and interaction terms.

A further aid in identifying the data and preparing them for regression runs was the concept of organizing the data files by Sectors.

The large number of principal, reading teacher, classroom teacher, school, and student variables could be collapsed rather neatly into five data files. The use of sectors was a great aid in preparing and interpreting the regression runs, especially in the early stages before many of the variables had been eliminated. Variables from each sector were present in every regression equation. The five sectors utilized were:

1. The Student/Socio-Economic Status Sector, which consisted of individual student characteristics of race, sex, family income, absences and achievement.
2. The School Characteristics Sector, which consisted of variables which described the school: its capacity, enrollment, racial composition, attendance, achievement, staff, and economic indicators.
3. The Teacher Characteristics Sector, which consisted of the educational background and experience of each teacher as well as the teacher's race, sex, teaching strategies, perceptions and attendance.
4. The Instructional Support Sector, which consisted of the combined Principal and Reading Teacher files, as well as selected Teacher variables that had to do with support of the reading program.
5. The Reading Program Sector, which consisted of those variables which actually described the approach and program as well as activities that directly related to the teaching of reading.

G. Specification of Reading Approaches

Since the study was primarily concerned with successful practices in reading, among the most critical variables were reading approaches and programs. Approaches were included as variables in each regression run. The approaches are Specific Skills, Traditional Basal, Linguistic Basal, and Linguistic Programmed. These approaches represent all 12 commercial programs used in the 94 fourth grade classrooms included in the study.

1. The Specific Skills Approach involves different sets of recommended materials at varied skill levels, tailored to the student's needs. Materials cover the areas of decoding, study skills, comprehension, and literature. Various publisher and teacher-made programs are used in an effort to supplement basic reading instruction in the classroom.

2. The Traditional Basal Approach utilizes a graded series of "readers" as its basic component, with many supplementary materials. Beginning with the development of a sight vocabulary, basals generally feature controlled vocabulary and skills with emphasis on comprehension. Basals are usually identified by the name of the publisher.

3. The Linguistic Basal Approach utilizes a graded series of "readers" as its basic component. The emphasis is jointly on phonics and comprehension. In these series, words are presented in groups based on sound-symbol relationships. Short vowels are usually presented first.

4. The Linguistic Programmed Approach is basically a linguistic approach with a strong decoding emphasis. The presentation of materials

is in small steps (frames). A response is required from the learner and immediate feedback is provided. Comprehension materials are used in conjunction with the program.

H. The Creation of The Experimental and Cross-Validation Files

Once the data were complete for all 1,828 fourth grade subjects, they had to be analyzed. Analysis should include cross-validation. Observational studies, usually due to their magnitude and expense, are difficult to cross-validate. Without some type of cross-validation procedure, however, the ultimate interpretation may be suspect.

The luxury of having so large a data base made cross-validation possible. Before a single regression equation was formulated, the final file was run with a program which randomly selected 25 percent of the cases.[6] These 465 subjects were placed on a separate tape. There they "rested" while an experimental file of 1,363 subjects was mined extensively.

Only after the final "best fit" equation had been run and recorded, was the cross-validation file accessed. This report represents the results of both the experimental sample (mined data) and the cross-validation sample (on which the "best fit" equation was tested).

[6]The program generated a random number for each case in succession for assignment to the experimental file or the validation file; thus every case had a 25 percent probability, independently of every other case, of being assigned to the validation file.

V. MODEL AND ESTIMATION PROCEDURES

In conducting research, the standard, scientific procedure is, of course, to (1) state the theory, (2) obtain the data appropriate to each theoretical concept involved, and (3) empirically test the theory, using standard tests of significance as the guideposts for measuring the success of the test. It is no accident, however, that, most commonly, writers of doctoral dissertations in economics and education obtain some data first; then examine the data (mine the data?), and then go back to write Chapter I on the theory! The fact is that in many areas of education, as in many areas of economics, there is no solid body of theory to draw upon, and the data are far from ideal in reflecting theoretical concepts. The standard scientific procedure, therefore, needs to be revised to accommodate the realities of the discipline.

A. General Description of Model and Estimation Procedures. In this study, which examines the determinants of reading achievement growth, there is no one agreed-upon body of theory to test. What has been done, then, in its absence, is to substitute an alternative way of arriving at a theoretical model and a procedure for testing it. More specifically, the following steps were taken:

1. The data in a sample of 1363 students were looked at to see what they said--i.e., through a series of multiple regression equations they were mined extensively in an experimental sample.
2. The final equation was regarded as The Theory--the hypothesized relationship between growth in reading achievement between the third and fourth grades and many inputs.

3. t-tests were run on the coefficients, but were used, of course, only as general guideposts, since the data were mined.
4. The Theory, the regression equation resulting from the mining, was tested on a randomly selected, totally unexplored sample of 465 students from the same population of students--the cross-validation sample.
5. Tests were run for examining the equality of the coefficients between the two samples.
6. The final conclusions and policy implications were drawn from the two sets of regression results, from the tests on the equality of the coefficients and from a liberal interpretation of the coefficients and the standard tests of significance.

The model which was developed was of the general form, $\Delta R = F(KID, PRIN, SCHL, TEACH, READ)$. That is, the combined wisdom of the literature on reading achievement and the personnel in the School District of Philadelphia responsible for designing and monitoring the reading programs, suggested that reading growth between the third and fourth grades (ΔR) is a critical educational junction; and that it is a function of many characertistics of the pupil (KID), characteristics of the school principal (PRIN), characteristics that described the atmosphere of the school (SCHL), characteristics of the student's teacher (TEACH), and characteristics of the reading program and the way in which it was delivered to the students (READ).

The relationship was examined using single equation multiple regression. In the experimental sample, dummy variables and other non-linearities were explored. Interactions of pupil achievement with other inputs were

examined extensively, consistent with the regularly stated theory that students of different abilities react differently to the characteristics of their schools, their peers, their teachers, and their skills programs. All the findings were checked and rechecked against alternative specifications for robustness--only those variables which remained strong throughout the mining process were retained.

B. Description of Table 1. The results of the exploration of the data in the experimental and cross-validation samples are presented in Table 1.

In Column (]) the variables are listed; where interaction terms are used (the interaction between student achievement and the Linguistic Basal reading approach, in lines [29,30] for example), three values for student achievement, covering the range of results recorded, were selected to expose the range of results.

In Column (2) the coefficients and t-statistics for the experimental sample are listed. Since the expression "experimental sample" is a euphemism for "worked-over" sample, the t-statistics should be read with the recognition that they are over-stated, reflecting the understatement of the standard errors. The t-statistics associated with the interaction terms (2.44 in the case of Linguistic Basal line [30], for example) only indicate, of course, whether or not different achievement levels do or do not make a difference in the impact of the input on reading achievement growth. The t-statistics associated with the interaction at specific values (+.28, +4.32, +3.83 in the case of Linguistic Basal, for example) indicate whether or not the input has a significant impact on students of that ability. This test was of the form:

$$\hat{\theta} = \frac{\hat{\alpha} + \hat{\beta}y^*}{\sqrt{\hat{\sigma}^2_{\hat{\alpha}} + \hat{\sigma}^2_{\hat{\beta}}y^{*2} + 2\hat{\sigma}_{\hat{\alpha}\hat{\beta}}y^*}},$$

where $\hat{\alpha}$ is the coefficient of the intercept term (Linguistic Basal line [29] for example), $\hat{\beta}$ is the coefficient of the interaction term (Linguistic Basal x Third Grade Score, line [30], for example), y^* is the specific value of taken by the interacting variable (Third Grade Score = 250, for example), $\hat{\sigma}^2_{\hat{\alpha}}$ is the variance of the intercept term, $\hat{\sigma}^2_{\hat{\beta}}$ is the variance of the interaction term, and $2\hat{\sigma}_{\hat{\alpha}\hat{\beta}}$ is the covariance.

In Column (3) the coefficients and t-statistics for the validation sample are listed. These results emerged from taking the variables listed in Column (2), the end-product of tortuous experimentation, and testing it on the 465 randomly selected virgin data. The t-statistics laid out for this sample are meaningful--they are consistent with the underlying assumptions of significance tests. It is the results of Column (3) which should receive the heaviest weight in the final interpretation of results.

The results of tests of the similarity of the coefficients of the two samples are shown in Column (4). Though all the $\tilde{\theta}$ are below 2, it is not clear how seriously they should be taken. The standard errors of the experimental sample, as has been indicated, are understated. In addition, the standard errors of the validation sample will be higher than the standard errors of the experimental sample, because the sizes of the samples differ. (Thus, the square root of the ratio of the sample sizes is 1.7 ($\sqrt{1363 \div 465}$), and the standard errors of the coefficients of the experimental sample vary around 1.7 times the standard errors of the coefficients of the validation sample.) The test for similarity used

for the variables without interaction took the form:

$$\mathcal{Y} = \frac{\hat{\beta}_1 + \hat{\beta}_2}{\sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}},$$

where $\hat{\beta}_1$ is the coefficient of the experimental sample, $\hat{\beta}_2$ is the matching coefficient of the validation sample, $\hat{\delta}_1$ is the standard error of the coefficient of the experimental sample, and $\hat{\delta}_2$ the standard error of the coefficient of the validation sample. For variables with interactions, the test was of the form:

$$\mathcal{Y} = \frac{(\hat{\alpha}_1 + \hat{\beta}_1 y^*) - (\hat{\alpha}_2 + \hat{\beta}_2 y^*)}{\sqrt{[\hat{\sigma}^2_{\hat{\alpha}_1} + \hat{\sigma}^2_{\hat{\beta}_1} y^{*2} + 2\mathrm{cov}(\hat{\alpha}_1\hat{\beta}_1)] + [\hat{\sigma}^2_{\hat{\alpha}_2} + \hat{\sigma}^2_{\hat{\beta}_2} y^* + 2\mathrm{cov}(\hat{\alpha}_2\hat{\beta}_2)]}},$$

where the terms with 1 as a subscript refer to the experimental sample, and the terms with 2 as a subscript refer to the validation sample. A pragmatic regression strategy calls for using all these results in the final interpretations.

VI. WHAT WORKS IN READING?

Many legally implementable results emerged from the extensive empirical analysis which should be useful to educational policy makers in the School District of Philadelphia, and suggestive to other school districts.

A. Impact of Pupil Characteristics. Which characteristics of students over which the School District has some control, can contribute to fourth grade reading growth? Attending school[7] (line 6) clearly increases the probability that the student will have a larger increase in reading scores between the third and fourth grades than those who are present fewer days. With somewhat less conviction--because the robustness of the coefficients throughout the exploration of the experimental sample, and passing the Column (4) test is somewhat offset by the results in Column (3)--one can say that attending kindergarten (line [4]), and having fewer unexcused absences (line [7]) contribute to fourth grade reading growth. It is clear that the higher a student's third grade reading score (line [3]), the less the fourth grade reading growth. Apart from the standard regression-toward-the-mean explanation, there is a general view that there is a fourth-grade slump for the students that were whizzing along through the first three years.

Equally important to note, are characteristics of pupils which were explored and found to be without significance: the pupil's family income, whether or not a pupils was bussed to relieve overcrowding, the number of residences the pupil had while in school, and whether or not the pupil attended kindergarten in the same school as his or her fourth grade.

---

[7] A good case can be made for treating attending school [line (6)] and unexcused absences [line (7)] as dependent variables. Reverse causality seems somewhat tenuous, however, since elementary school children are usually unaware of their standardized test scores. When the equations were estimated without these variables, the remaining coefficients were not changed in a way which altered the conclusions.

B. Impact of Principal Characteristics. One of the more interesting results of this study is the evidence it provides on the impact of a principal's involvement in reading on reading achievement. Principals who were reading professionals had fourth grade students who achieved more reading growth (line [8]). It has been suggested that the explanation lies in the fact that such a principal is likely to be an active instructional leader in the area. Because he or she is conversant with the subject, reading and classroom teachers and the principal "speak the same language", and students benefit from their exchange. In addition (line [9]), the evidence of this study is that the more a principal observes the fourth grade reading classes, the greater the development of the students. The message from these samples, at least, is clear--involvement of elementary principals in reading has a payoff in reading achievement growth.

Many other characteristics of principals were examined and found to be without significance: whether or not the principal had previous administrative experience, the amount of consultation received from reading program publishers and the District office, whether or not the principal had an elementary school teaching certificate, his or her highest degree, the number of years he or she was a classroom teacher, the number of years in the school, the number of years as a principal, and the perceptions of the principal about home support and teacher expertise.

C. Impact of School Characteristics. The combined evidence of the results indicates strongly that students in schools with more supportive staff per pupil (line [18]) and with fewer non-white students (line [19]) are more likely to achieve greater fourth grade reading growth. Clearly, here as elsewhere, direct cause and effect conclusions are inappropriate. We do not know whether the fact that there are more supportive personnel per pupil is of direct help to students, or whether the characteristic

of the school which resulted in a higher ratio is the helping agent. Similarly, the relationship between the racial mix of the school and fourth grade reading growth is more likely explainable by a set of affective characteristics associated with that mix which is not picked up by the other variables.

A slightly less strong finding but a very robust one throughout the exploration of the experimental sample, is the negative impact of being in a K-6, school (line 12), compared with the alternatives. A K-5 organization compared with K-8 makes no difference; K-7 seems to be superior to K-8. The evidence of this study suggests that, in terms of impact on reading growth, K-7 is best, K-5 and K-8 are about the same, and K-6 is distinctly inferior.

The results also indicate fairly strongly that students in schools with more high achievers (lines [16,17]) do better, though the very high achieving students are not affected by this aspect of the school climate. Less strong, but not to be ignored, is the finding that fourth grade reading growth is negatively associated with being in a school with more very new teachers (lines [14,15]); and that it is the students at grade level and above who are most affected; it seems to be irrelevant to the ones distinctly below grade level.

There is some slight suggestion that older schools are better (line [10]); but the appropriate conclusion from the burden of evidence is to ignore the suggestion. There are several reasons: (1) the coefficient in the experimental sample is relatively weak, (2) its insignificance in the cross-validation sample is unambiguous; and (3) reasonable priors would lead one to assume that there are characteristics other than age itself

which must have produced such a result.

Other characteristics of the schools were examined and found to be not significant: median income of the school feeder areas, proportion of students from familes on AFDC, proportion of students below the 16th percentile, proportion of non-white professional staff, and mobility of the school population.

D. Impact of Teacher Characteristics. The qualities of a teacher which contribute to successful results for fourth grade readers must be many, only a few of which are readily quantifiable. Of these, the findings indicate strongly that the presence of the teacher (as opposed to some sort of substitute) is important: students who had teachers with more pay periods with no absences (line [20]) did distinctly better, as did students who had teachers who attended fewer outside conferences (line [21]). And students who had teachers who were teaching the fourth grade for the first time (line [23]) did distinctly worse.

There is less strong evidence that female teachers (line [22]), were associated with more fourth grade reading achievement, and a slight suggestion that male students do better with female teachers, and female students do better with male teachers (line [24]).

A large number of other characteristics of teachers were explored and found to be not significant: The Gourman rating of his or her undergraduate school, the number of graduate reading and language arts courses, the National Teachers' Examination score, race, the number of years as a classroom teacher, and whether or not this was the first time the teacher was teaching in the present school, whether he or she had requested assignment to the present school, or whether there was teacher involvement in the school reading plan.

E. Impact of Reading Instruction Characteristics. There seems to be compelling evidence that there are a number of characteristics of reading instruction which are stongly associated with fourth grade reading growth. The more minutes per week spent on individual-independent reading (line [25]) the better; if the classroom teacher uses fewer aides per week (line [26]), there is more growth in reading scores; and if the classroom teacher says he or she would select the particular reading program again (line [27]), the student is likely to show more reading improvement--a reflection of more teacher confidence, perhaps.

Most interesting of all, the Linguistic Basal approach produces better fourth grade reading growth for students at or above grade level than any of the three other approaches used in the Philadelphia School District--Specific Skills, Traditional Basal, or Programmed Linguistic. Probably no finding in the study was more robust. Many, many alternative specifications were examined--Linguistic Basal was always associated with higher rates of learning. Low achieving students were not helped more by one type of program than another.

The class size debate must continue on. In this study, the evidence is that students do better in larger classes (line [35]). Alternative specifications--nonlinear and interactive--were extensively explored. But the finding remains! Certainly, many studies point to the opposite finding, but others have found the same one. In interpreting the finding, however, it is important to emphasize that it is a finding which emerges when many other variables are controlled--that is, what the positive coefficients are saying is that larger classes are better, after controlling for such instructional characteristics as the degree of individualization

in teaching reading.

The "right" grouping for reading instruction is an area extensively discussed in the literature. The findings for fourth graders in the School District of Philadelphia (lines [31-34]) is that students taught reading in a class where individualized instruction only or small groups only are used are worse off than those taught where there is some large grouping in the room. Students in classes where a mixture of individual and small groups are used to teach reading do as well as those with some large groupings, and better than those in classes where there is only individual or only small grouping. One interpretation of these results is that, again, the active presence of the teacher plays a major role in learning. A hypothesis is that the reason the individual and small groupings are less effective is that the teacher can then only be involved with a few students at a time, while, in the large groupings, the teacher is actively involved with more students more of the time.

Finally, there was a strong suggestion in the experimental sample that certain positive things happen when a program is used for the first time (line [28]). It has been suggested by reading specialists that the particular Hawthorne-effect advantage here is that teachers may pay much stricter attention to the instructions associated with a reading program during the first year. The cross-validation sample did not support this hypothesis, however.

There were many other characteristics of reading instruction explored, but found not to be significant. Included were the number of minutes per day of reading instruction, the time of day reading was taught, the number of years the reading teacher was a classroom teacher, the number of reading

and language arts courses taken by the reading teacher, the generally used professional development methods (outside visits to reading programs, seminars attended, professional organization participation), consultation from the District office and reading program publisher, and the specific way in which the reading teacher helps the classroom teacher (helps structure the classroom, demonstrates techniques, provides professional information). It should be noted that this study was conducted during the fifth year of the District Reading Programs. Some of the things found to be of no importance at that point might have been of critical importance during the first year or two of the program.

The conclusion that emerges is that the choice of reading approach makes a difference to all but the very low achievers, and that the active, direct efforts and time of the teacher (not substitutes for the teacher) makes a significant difference and is reflected on several inputs.

F. Summary

The findings of most immediate interest to policymakers are those over which they have some control. What does this study contribute to the issues that educators are concerned with?

The issue explored in the study is fourth grade reading achievement growth as measured by changes in standardized test scores. It is not likely that fourth graders are completely different from all other grades, but one should be careful in concluding that all of the findings would be exactly the same for first graders or twelfth graders.

There are many studies, extensively developed in the literature, which document the problems in any precise reading of the statistical results of this type of education production function investigation. Thus, the intent here is to interpret the findings in a broad, directional

manner--that is, to indicate the positive or negative effect of only the robust findings, not to suggest that we know precisely how much reading scores can be altered by altering an input by a specified amount. While the caveats are being laid out, it is also important to keep in mind that the study has been done only with students in the School District of Philadelphia. It has been shown from this study, better than from previous studies, what is effective with them, because of the extensiveness of the classroom, program, and pupil-specific data, and because of the regression strategy used to analyze the data. What is yet to be determined, however, is how eternal the revealed verities are. It remains for others to test the hypotheses in other school districts. It remains for the School District of Philadelphia to be ever-ready to test out the effects of new techniques and organizational changes--programs involving foreign languages, multi-media, and parent participation; as well as any changes in policy or practice resulting from findings generated by this study.

This study identified certain combinations of elements that made a difference in reading achievement. There are limitations in the interpretation of all studies. This study, for example, could not examine all possible combinations of variables. Nor did the study tell why certain variables made a difference. To try to find out why, additional studies of a different kind would be needed.

VII. POLICY IMPLICATIONS

The study has yielded considerable information about reading that top management and key decision makers of the School District of Philadelphia might find both interesting and useful. The following would seem to be most important implications for policy.

A. Spending more days in school is associated with learning--probably because it is a signal of pupil motivation, and because things happen in school which make being there important to learning. For either reason, attendance records are important measures of the success of programs and how they are implemented.

B. Neither frequency of residential moves nor bussing undertaken to relieve overcrowding is associated with poorer learning. There may be other detrimental effects from those factors, but the evidence does not point to effects on fourth grade reading.

C. Involvement of elementary school principals in reading, through previous professional training in reading and current observation, is associated with greater learning. If fourth grade reading growth is important then there is a case for paying attention to these characteristics of principals--rather than to characteristics such as degrees, experience, consultation, sex or race, which are not visibly associated with learning.

D. Organizing elementary schools in a K-6 format, compared with the alternatives, is associated with poorer learning; the K-7 seems most effective for fourth grade reading development. Schools with more supportive staff per pupil are similarly effective. Frequently additional supportive staff is funded through federal sources, which suggests that this money is well targeted.

E. The composition of the student body is associated with learning--low and middle achieving students do better in schools with more high achievers. Translating this finding into policy, however, seems inappropriate, since in this area, the contribution that changes might make to fourth grade reading is clearly only a very small portion of what the policy objectives are.

F. The median income of the school feeder area, the proportion of students from families on AFDC, and the proportion of students, below the 16th percentile are all not accurate flags for predicting fourth grade learning. The policy implication of this finding is important. It suggests that allocating federal funds on the basis of school-wide low income and low achievement characteristics is not appropriate, if the object is to help the disadvantaged student. The targeting has to be pupil-specific.

G. More fourth grade reading growth is associated with more direct contact of student and teacher--with fewer personal or professional absences that are taken from teaching time. This suggests that educational procedures and rewards which increase the contact are likely to be beneficial to learning, and that professional training may be best accomplished outside the hours of a standard teaching year. Very new teachers, it would appear, operate to the disadvantage of the fourth grade reader--perhaps suggesting the need for some experienced reading support to brand new teachers.

H. The oft-repeated findings that the characteristics of teachers which are explicitly rewarded--experience, extra credits of education, and degrees--are not associated with learning, is further confirmed in this study. If student learning is the main objective of a school system, then some rewards based on achieving learning warrant consideration.

1. The policy implications from these findings on different characteristics of reading instruction are clear: the more time spent on independent reading, the more learning; the enthusiasm of the teacher for the program is beneficial; the Linguistic Basal approach is associated with more learning growth than any of the alternatives currently used in Philadelphia; and when the teacher has more direct contact in teaching reading (uses few aides and deals with more students simultaneously--leaving fewer unattended), students learn more. Many of the standard approaches to the training and development of reading teachers did not show evidence of being associated with more learning for the students.

Beyond these specific policy implications, there are some of a more general nature. First, the findings suggest that some reallocation on inputs can occur which would use no additional resources, but which would improve the reading achievement growth of fourth graders. The findings suggest, for example, that money would be better spent on keeping the teacher with the student than on the usual staff development activities which take the teacher away.

Second, this type of rigorous evaluation procedure is a useful method for school districts to use in looking at their educational processes. The technology described in this study is used in many other disciplines--sociology, history, psychology; but it has seldom been used in educational inquiry, and rarely been used by individual school districts. It is a technology which can help to substitute for the price system discipline available to the private sector. System analysis is the public sector's efficiency tool.

We are in a new era of scarcity and restraint in which the difficulties of achieving efficiency in the public sector will not be papered over by constantly rising levels of prosperity. Genuine appraisal of aspects of the educational process need a combination of general wisdom and modern assessment technology.

TABLE 1

REGRESSION RESULTS FOR FOURTH-GRADE SAMPLES OF PHILADELPHIA SCHOOL DISTRICT STUDENTS, 1975 - 1976

(1) Variables Representing Characteristics of:	(2) Experimental Sample		(3) Cross-Validation Sample		(4) Test of Equality of Coefficients
	coeff	t-ratio	coeff	t-ratio	
Students					
(1) Income	-.001	(-.70)	.001	(.56)	-.84
(2) Race	-18.39	(-4.27)	-19.50	(-2.47)	.12
(3) Third Grade Score	-.42	(-4.53)	-.47	(-3.00)	.26
(4) Kindergarten	5.73	(1.82)	-3.15	(-.56)	1.39
(5) Sex	12.45	(2.73)	8.90	(1.14)	.39
(6) Days Present	.34	(2.16)	.99	(3.42)	-1.95
(7) Unexcused Absences	-.42	(-1.58)	.58	(1.36)	-1.99
Principals					
(8) Reading Experience	16.11	(3.91)	13.68	(2.04)	.31
(9) Observes Reading	.71	(5.05)	.71	(2.84)	.001
Schools					
(10) Age	.12	(1.33)	.04	(.24)	.47
(11) K-5 organization	3.22	(-.72)	7.24	(.99)	-1.22
(12) K-6 organization	-6.40	(-2.64)	-6.56	(-1.58)	.03
(13) K-7 organization	10.31	(2.31)	4.80	(.71)	.68
(14) % Teachers < 2 years	97.06	(1.71)	89.86	(1.02)	
(15) (14) x (3)	-.38	(-2.52)	-.26	(-1.08)	
(3) = 250	2.06	(.13)	24.86	(.74)	-.18
(3) = 400	-54.94	(-3.23)	-14.14	(-.55)	-.26
(3) = 550	-111.94	(-3.46)	-53.14	(-1.02)	-.31
(16) % Students > 84th percentile	272.28	(3.98)	199.94	(1.71)	
(17) (16) x (3)	-.44	(-2.56)	-.32	(-1.07)	
(3) = 250	162.28	(5.46)	119.94	(2.40)	.27
(3) = 400	96.28	(5.01)	71.94	(2.15)	-.12
(3) = 550	30.28	(.90)	23.94	(.40)	.03
(18) Supportive Staff per pupil	.97	(4.39)	.89	(2.36)	.17
(19) % Non-white students	-15.67	(-3.28)	-23.53	(-2.71)	.79
Teachers					
(20) Pay Periods with no absence	.39	(2.68)	.86	(1.36)	.17
(21) Attends outside Conferences	-12.96	(-3.28)	-19.03	(-3.01)	.81
(22) Sex	17.61	(3.40)	8.90	(1.15)	1.66
(23) First time teaching grade 4	-22.31	(-4.78)	-10.04	(-1.28)	-1.34
(24) Teacher Sex = Pupil Sex	-8.22	(-1.80)	2.70	(.35)	-1.23

TABLE 1 (Continued)

REGRESSION RESULTS FOR FOURTH-GRADE SAMPLES OF PHILADELPHIA SCHOOL DISTRICT STUDENTS, 1975 - 1976

(1) Variables Representing Characteristics of:	(2) Experimental Sample		(3) Cross-Validation Sample		(4) Test of Equality of Coefficients
	coeff	t-ratio	coeff	t-ratio	
Reading Instruction					
(25) Minutes Independent Reading	.11	(4.57)	.12	(2.93)	-.21
(26) Aides per Week	-1.25	(-1.64)	-1.36	(-1.03)	.07
(27) Select Program Again	7.59	(2.42)	10.26	(1.92)	.43
(28) First Year with Program	20.73	(5.63)	-10.04	(-1.28)	1.45
(29) Linguistic Basal Approach	-42.42	(-1.50)	-110.40	(-2.28)	
(30) (29) x (3)	.18	(2.44)	.36	(2.80)	
(3) = 250	2.58	(.28)	-20.40	(-1.14)	.36
(3) = 400	29.58	(4.33)	33.60	(2.87)	-.03
(3) = 550	56.58	(3.83)	87.60	(3.39)	-.29
(31) Individualizes Only	-67.18	(-2.45)	-40.14	(-.89)	
(32) (31) x (3)	.17	(2.43)	.06	(.55)	
(3) = 250	-24.68	(-2.26)	-25.14	(-1.32)	-.001
(3) = 400	.82	(.08)	-16.14	(-1.50)	.22
(3) = 550	26.32	(2.01)	-7.14	(-.31)	.36
(33) Small Groups Only	.15	(.03)	-17.52	(-1.89)	1.67
(34) Individual and Small Groups	9.71	(1.60)	-9.17	(-.85)	1.53
(35) Class Size	1.27	(2.71)	1.91	(2.42)	-.70
Constant	-9.02		10.84		
Adjusted R^2	.23		.32		
N	1,363		465		

Dependent Variable: Fourth Grade Scale Score (February 1976) Minus Third Grade Scale Score (February 1975) on the Reading Comprehension Subtest of the California Achievement Test, 1970 edition.

Note: The means, standard deviations and units associated with each variable are presented in Table 2.

MEANS AND STANDARD DEVIATIONS OF VARIABLES LISTED IN TABLE 1

Variables Representing Characteristics of:	Units	Experimental Sample (N=1363) Mean	S.D.	Validation Sample (N=465) Mean	S.D.
Students					
(1) Income	dollars	9036.52	3159.00	8797.70	3159.02
(2) Race	white, non-white	.65	.48	.67	.47
(3) Third Grade Score	scale scores	385.06	67.74	385.33	68.39
(4) Kindergarten	yes, no	1.79	.41	1.81	.39
(5) Sex	male, female	1.51	.50	1.51	.50
(6) Days Present	number	130.26	10.41	131.51	9.13
(7) Unexcused Absences	number	2.49	6.44	2.53	6.30
Principals					
(8) Reading Experience	yes, no	1.34	.47	1.31	.46
(9) Observes Reading Classes	number	4.78	11.98	4.95	12.11
Schools					
(10) Age	years	31.66	25.19	32.72	26.02
(11) K-5 Organization	yes, no	.03	.49	-.02	.51
(12) K-6 Organization	yes, no	.55	.68	.46	.72
(13) K-7 Organization	yes, no	-.01	.46	.01	.53
(14) % Teachers < 2 years	%	.19	.13	.21	.16
(16) % Students > 84th percentile	%	.20	.13	.19	.13
(18) Supportive Staff per pupil	#/thousand	29.57	11.01	29.56	11.03
(19) % Non-white Students	%	.50	.45	.44	.45
Teachers					
(20) Pay periods with no Absences	number	13.92	3.79	14.22	3.82
(21) Attends outside Conferences	yes, no	1.19	.39	1.21	.41
(22) Sex	male, female	1.92	.28	1.91	.29
(23) First time teaching Grade 4	yes, no	1.14	.35	1.13	.34
(24) Teacher Sex = Pupil Sex	yes, no	1.51	.50	1.51	.50
Reading Instruction					
(25) Minutes independent reading	minutes	73.90	60.31	75.23	61.19
(26) Aides per Week	number	2.29	2.31	2.44	2.33
(27) Select Program again	yes, no	1.56	.50	1.56	.50
(28) First Year with Program	yes, no	1.24	.43	1.20	.40
(29) Linguistic Basal Approach	yes, no	.09	.29	.10	.31
(31) individualizes Only	yes, no	1.11	.32	1.10	.30
(33) Small Groups Only	yes, no	1.65	.48	1.65	.48
(34) Individual and Small Groups	yes, no	1.11	.32	1.14	.34
(35) Class Size	number	30.12	3.49	30.10	3.45
Dependent Variable					
Gain from 1975 to 1976 CAT-70 Reading Comprehension	scale scores	27.44	52.50	31.33	54.87

REFERENCES

Alexander, K. L., and McDill, E. L. "Selection and Allocation Within Schools: Some Causes and Consequences of Curriculum Placement." *American Sociological Review*, 1976, 41.

Averch, H. A., et al. *How Effective is Schooling: A Critical Review and Synthesis of Research Findings*. Santa Monica, Calif.: Rand, 1972.

Barr, R., and Dreeben, R. "Instruction in Classrooms. In Lee S. Shulman, (ed.). *Review of Research in Education 5*, Itasca, Ill.: Peacock, 1977.

Burstein, L. "The Role of Levels of Analysis in the Specification of Educational Effects." Paper commissioned by the Educational Finance and Productivity Center, University of Chicago, 1978.

Chall, J. S., *Learning to Read: The Great Debate*. New York: McGraw-Hill, 1967.

Coleman, J. S., et al. *Equality of Educational Opportunity*. Washington: Government Printing Office, 1966.

Cooley, W. W. "Explanatory Observational Studies." Pittsburgh: University of Pittsburgh, Learning Research and Development Center, 1978.

Della-Piana, G. M., and Endo, G. T. "Reading Research." In R.M.W. Travers (ed.), *Second Handbook of Research on Teaching*. Chicago: Rand McNally, 1973.

Edmonds, R. R. and Federiksen, J. R. Draft Article, "Search for Effective Schools: The Identification of Schools that are Instructionally Effective for Poor Children." Cambridge, Mass., 1977.

Guthrie, J. T., et al. "A Study of the Locus and Nature of Reading Problems in the Elementary School." International Reading Association. Newark, Delaware, 1976.

Hanushek, E. *Education and Race*, Lexington, Mass.: D.C. Heath, 1972.

_____. "The Production of Education, Teacher Quality, and Efficiency," in *Do Teachers Make a Difference?*, Office of Education, Department of Health, Education and Welfare, Bureau of Educational Personnel Development, Washington, 1970.

Jencks, C. et al. *Inequality: A Reassessment of the Effect of Family and Schooling in America*. New York: Basic Books, 1972.

Jensen, A. R. "How Much Can We Boost I.Q. and Scholastic Achievement?" *Harvard Educational Review*. Winter, 1969.

Murnane, R. J. *The Impact of School Resources on the Learning of Inner City Children*. Cambridge, 1975.

Spady, W. G. "The Impact of School Resources on Students." In F. N. Kerlinger (ed.), *Review of Research in Education*. Itasca, Ill.: Peacock, 1973.

Summers, A. A., and Wolfe, B. L. *Equality of Educational Opportunity Quantified: A Production Function Approach*. Unpublished manuscript, Federal Reserve Bank of Philadelphia, 1974.

_____, and _____. "Which School Resources Help Learning? Efficiency and Equity in Philadelphia Public Schools." *Federal Reserve Bank of Philadelphia Business Review*. February, 1975.

_____, and _____. "Do Schools Make a Difference?" *American Economic Review*. September, 1977.

Winkler, D. R. "Educational Achievement and School Peer Composition." *Journal of Human Resources*. Spring, 1975.

SOURCES OF DATA

Variables Representing Characteristics of:	Source
Students	
(1) Income	Federal Reserve Bank block income number
(2) Race	School District of Philadelphia Pupil Directory Information File (PDIF)
(3) Third Grade Score	Student record card
(4) Kindergarten	Student record card
(5) Sex	PDIF
(6) Days Present	Student roll sheet
(7) Unexcused Absences	Student roll sheet
Principals	
(8) Reading Experience	Principal interview
(9) Observes Reading Classes	Teacher interview
Schools	
(10) Age	Superintendent's Management Information Center (M.I.C.)
(11) K-5 organization	M.I.C.
(12) K-6 organization	M.I.C.
(13) K-7 organization	M.I.C.
(14) % Teachers< 2 years	M.I.C.
(16) % Students> 84%ile	M.I.C.
(18) Supportive Staff per pupil	M.I.C. $\frac{\text{full time staff-classroom teachers}}{\text{avg. on roll}} \times 1{,}000$
(10) % Non-white students	M.I.C.
Teachers	
(20) Pay periods with no absences	Office of Personnel and Labor Relations
(21) Attends outside conferences	Teacher interview
(22) Sex	Teacher interview
(23) First time teaching Grade 4	Teacher interview
(24) Teacher Sex=Pupil Sex	Teacher Interview and P.D.I.F.
Reading Instruction	
(25) Minutes independent reading	Teacher interview
(26) Aides per week	Aide interview and Teacher interview
(27) Select Program again	Teacher interview
(28) First Year with Program	Teacher interview
(29) Linguistic Basal Approach	Teacher interview
(31) Individualizes only	Teacher interview
(33) Small Groups only	Teacher interview
(34) Individual and Small Groups	Teacher interview
(35) Class Size	Class Roll Book and Teacher Interview
Dependent Variable	
Gain from 1975 to 1976 CAT-70 Reading Comprehension	Student Record Card

DIGEST OF RESULTS

After the data mining and the cross-validation, it is clear that certain legally controllable variables contribute to gains in a pupil's reading achievement between the third and fourth grade. Many other variables were included in the study during its "mining" phase. However, they were not included in the final equation because, despite the original hypotheses, they made no significant difference in reading achievement gain.

Variables that contribute to achievement gain, and those that make no difference are listed by sector.

Student/Socio-Economic Status

These variables contribute to achievement gain:

- attending school regularly
- having few unexcused absences
- going to Kindergarten

These variables make no significant difference:

- whether the child was bussed to relieve overcrowding
- the number of times the child was late for school
- whether the child had Kindergarten in the same school
- the number of years and months the child has attended the school

School Characteristics

These variables contribute to achievement gain:

- being in a school with more supportive staff
- being in a school with more high achievers
- being in a school with fewer new teachers
- attending a K-7 school

These variables make no significant difference:

- the school average number on roll
- the school capacity
- the median income of the school feeder area
- the percentage of students in the school from families receiving AFDC
- the number of students retained in grade by the school
- the number of classroom teachers in the school
- the school's average daily attendance
- the percent of non-white professional staff in the school
- the percent of school mobility
- whether or not the school's classroom structure is "open" or "traditional"

Teacher Characteristics

These variables contribute to achievement gain:

- having a classroom teacher who attends regularly
- having a classroom teacher who attends fewer outside conferences
- having an experienced fourth grade teacher

These variables make no significant difference:

- the Gourman rating of the classroom teacher's undergraduate college or university
- the classroom teacher's score on the National Teachers Examination
- teacher attendance at staff development sessions
- the race of the teacher
- whether or not the teacher requested assignment to the school
- whether or not the teacher had direct involvement in the school reading plan
- the number of years the teacher has been teaching in the system
- the number of years the teacher has been teaching in the school
- the number of graduate courses in reading and/or language arts the teacher had taken

Instructional Support

These variables contribute to achievement gain:

- having a principal with a background in reading
- having a principal who frequently observes reading classes

These variables make no significant difference:

- whether or not the principal had previous administrative experience
- consultant service from reading program publishers and district office staff
- whether or not the elementary principal has an elementary teaching certificate
- whether or not the principal has a Masters, Masters +30 credits, or a Doctorate
- the race of the principal
- the sex of the principal
- whether or not the principal visits reading programs outside the school
- the principal's years of experience as a classroom teacher
- the principal's years of experience as a principal
- the number of staff development sessions at the district office the reading teacher attended
- the number of graduate courses in reading and/or language arts the reading teacher had taken
- whether or not the reading teacher demonstrated instructional techniques, helped structure classrooms or provided professional information to teachers
- the number of hours per week the reading teacher spends teaching students
- the race of the reading teacher
- the sex of the read ng teacher
- whether or not the reading teacher visits reading programs outside the school
- whether or not the reading teacher visits classrooms only at the request of the classroom teacher or principal

- number of professional contacts the classroom teacher has with the reading teacher per month
- whether or not the teacher receives feedback from the principal after observations
- whether or not the teacher perceives the reading teacher as being available when needed

Reading Program

These variables contribute to achievement gain:

- spending time reading independently
- having more direct contact with the classroom teacher
- having a classroom teacher who has confidence in his or her reading materials
- using a Linguistic Basal reading approach
- being in a class for reading which utilizes both large and small groups
- having classroom reading aide support fewer hours per week

These variables make no significant difference:

- whether or not the student is cycled for reading instruction
- whether or not there is additional instruction available for students experiencing reading difficulties
- whether the child has reading in the morning or afternoon
- the way in which the teacher makes use of the school library or the classroom library
- the minutes of reading per day
- whether or not the teacher modifies the reading program
- the number of minutes per week the teacher reads to the class
- whether or not the student uses a Programmed Lingusitic, Specific Skills or Traditional Basal reading approach